PB89-236459

Proceedings of the Workshop on One- versus Two-Generation Reproductive Effects Studies

(U.S.) Environmental Protection Agency, Washington, DC

Nov 88

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completi

1. REPORT NO. EPA/600/J-88/348	2.	3. PB89-236459
4. TITLE AND SUBTITLE Proceedings of the Workshop on One- vs Two-Generation Reproductive Effects Studies		5. REPORT DATE November 1988 6. PERFORMING ORGANIZATION CODE EPA/600/021
7. AUTHOR(S) Elaine Z. Francis & Gary L. Kimmel		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Office of Health & Environmental Assessment (RD-689) Human Health Assessment Group U.S. Environmental Protection Agency Washington, DC 20460		10. PROGRAM ELEMENT NO. 11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS		13. TYPE OF REPORT AND PERIOD COVERED Journal Article 14. SPONSORING AGENCY CODE EPA/600/21

15. SUPPLEMENTARY NOTES
Proceedings of Workshop held October 21-22, 1987, Arlington, Va. Published in Journal of the American College of Toxicology, Vol.7, No. 7

16. ABSTRACT

The Workshop on One- vs Two-Generation Reproductive Effects Studies was held on October 21-22, 1987. It was organized by the U.S. Environmental Protection Agency's Office of Pesticides and Toxic Substances and Office of Research and Development and was supported by the Agency's Risk Assessment Forum. The purpose of the workshop was to address the central question: Is a single-generation reproductive effects study sufficient to assess the reproductive toxicity potential of chemicals that do not bioaccumulate? In response to this question, the panel of nine participants concluded that, by itself, a one-generation reproductive effects study is insufficient to identify all potential reproductive toxicants and that a two-generation study is needed for an adequate assessment. The panel did not support placing much importance on bioaccumulation as related to reproductive toxicity testing mainly because it is not the only critical factor that may account for effects in a second generation but not in the first. In studying broader issues, the panel accomplished several goals that hopefully may serve to direct future development in reproductive toxicity testing: 1) the purpose of a reproductive effects study was defined, 2) a minimal set of end points necessary for adequate evaluation of reproductive toxicity was determined, 3) some alternative reproductive effects test methods were recommended, 4) greater flexibility in choosing a test protocol on a case-by-case basis was encouraged, and 5) areas that need further research were identified.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Processing Pages	Only.	

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

18. DISTRIBUTION STATEMENT Release to the public	19. SECURITY CLASS *(This Report)* Unclassified 20. SECURITY CLASS *(This page)* Unclassified	21. NO. OF PAGES 17 22. PRICE

JOURNAL OF THE AMERICAN COLLEGE OF TOXICOLOGY
Volume 7, Number 7, 1988
Mary Ann Liebert, Inc., Publishers

PB 89-236459

Proceedings of the Workshop on One- vs Two-Generation Reproductive Effects Studies[1]

ELAINE Z. FRANCIS[2] and GARY L. KIMMEL,[3] Co-Chairpersons

Participants:

Laurence D. Chitlik, EPA, Office of Pesticides and Toxic Substances
Mildred S. Christian, Argus Research Laboratories, Inc.
Thomas F. X. Collins, Food and Drug Administration
James C. Lamb, EPA, Office of Pesticides and Toxic Substances[4]
Anthony K. Palmer, Huntingdon Research Centre, UK
Joseph A. Ruddick, Environmental Health Directorate, Canada
Harold Zenick, EPA, Office of Research and Development

ABSTRACT

The Workshop on One- vs Two-Generation Reproductive Effects Studies was held on October 21–22, 1987. It was organized by the USEPA's Office of Pesticides and Toxic Substances and Office of Research and Development and was supported by the Agency's Risk Assessment Forum. The purpose of the workshop was to address the central question: Is a single-generation reproductive effects study sufficient to assess the reproductive toxicity potential of chemicals that do not bioaccumulate? In response to this question, the panel of nine participants concluded that, by itself, a one-generation reproductive effects study is insufficient to identify all potential reproductive toxicants and that a two-generation study is needed for an adequate assessment. The panel did not support placing much importance on bioaccumulation as related to reproductive toxicity testing mainly because it is not the only critical factor that may account for effects in a second generation but not in the first. In studying broader issues the panel accomplished several goals that hopefully may serve to direct future development in reproductive toxicity testing: (1) the purpose of a reproductive effects study was defined, (2) a minimal set of end points necessary for adequate evaluation of

[1]October 21–22, 1987, Sheraton National Hotel, Arlington, Virginia. Sponsored by U.S. Environmental Protection Agency Risk Assessment Forum.

[2]Office of Pesticides and Toxic Substances. *Present address:* Office of Research and Development.

[3]Office of Research and Development.

[4]*Present address:* Jellinek, Schwartz, Connolly, and Freshmen, Inc.

The views in this paper are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

reproductive toxicity was determined, (3) some alternative reproductive effects test methods were recommended, (4) greater flexibility in choosing a test protocol on a case by case basis was encouraged, and (5) areas that need further research were identified.

INTRODUCTION

THE WORKSHOP ON ONE- VS TWO-GENERATION REPRODUCTIVE EFFECTS STUDIES was held at the Sheraton National Hotel, Arlington, Virginia, October 21–22, 1987. The workshop was organized by the U.S. Environmental Protection Agency's Office of Pesticides and Toxic Substances and Office of Research and Development and was supported by the Agency's Risk Assessment Forum.

The purpose of the workshop was to address the question: Is a single-generation reproductive effects study sufficient to assess the reproductive toxicity potential of chemicals that do not bioaccumulate? The basis for raising this issue for discussion was a review of multigeneration studies by Christian,[1] which was commissioned by the Chemical Manufacturers Association. One of the conclusions of the review was: "In the absence of evidence of bioaccumulation, the studies reviewed provide limited evidence that a one-generation reproduction study is a sufficient test for determining the no-effect level for reproductive toxicity."[1] On the other hand, the review also stated "... when evidence of bioaccumulation is observed in the reproduction study or in chronic or biochemical studies, evaluation of more than one generation is indicated."[1]

The policy of the U.S. Environmental Protection Agency's (EPA) Office of Toxic Substances (OTS) and Office of Pesticides Program (OPP) (which together constitute the Office of Pesticides and Toxic Substances) and the Food and Drug Administration (FDA) has been to require at least a two-generation reproductive effects study to assess the potential reproductive toxicity of a chemical. Following Christian's report,[1] the chemical industry began submitting results of one-generation studies to OTS and contended that these were sufficient to address any concerns for reproductive toxicity.

OTS made an interim decision not to alter its long-standing policy of requiring a two-generation reproductive effects study. However, OTS recognized the need to explore this issue in greater detail and sought the assistance of EPA's Risk Assessment Forum to bring together a small group of experts with backgrounds in multigeneration testing. The workshop consisted of nine participants (Panel), with experience in either conducting or evaluating multigeneration reproductive effects studies. The number of participants was intentionally small to provide an environment that would facilitate discussion and best use the experience of the participants. The objective of the workshop was to review the hypothesis that a single-generation reproductive effects study is sufficient for chemicals that do not bioaccumulate. Christian's review[1] served as a starting point for the workshop, but the goal of the organizers was to look at some broader issues as well. Before the workshop, each participant was asked to respond to four issues: (1) the purpose of a reproductive effects study, (2) additional information that a two-generation study provides over a one-generation study, (3) bioaccumulation as a consideration in reproductive toxicity testing, and (4) alternative study designs. In addition, the participants were asked to suggest specific areas that should be addressed in the future. These issues and the participants' initial response served as the basis for the 2 days of discussion. The summary of that discussion follows.

ISSUE: PURPOSE OF A REPRODUCTIVE EFFECTS STUDY

The first issue addressed the purpose of a reproductive effects study and established a framework for discussion of the other issues. The participants noted that the basic design of a reproductive effects study is not uniform worldwide and that the purpose for gathering the data

is not always clear. Without some agreement about the expected end points to be assessed in such studies, discussion of specific design and evaluation issues is of limited use.

The participants supported the view that, from a risk assessment perspective, the primary purpose is to identify a reproductive hazard relative to other hazards and determine the levels of exposure in experimental animals that produce toxicity. From this and other information, the Agency may establish an allowable level of exposure for humans. For the purpose of risk assessment, the focus of a reproductive toxicity study is to identify reproductive effects in the context of other effects, which may or may not be related to reproduction. Thus, there is a bias toward emphasizing those end points that measure effects on the reproductive system. Nevertheless, other types of toxicity should not be ignored, and these data help form the context in which the reproductive effects are placed. In many instances of assessment, it is difficult to separate reproductive effects from other toxic effects. In any case, either effect is an important indicator that exposure can lead to an undesirable outcome.

The reproductive process involves a large number of organ systems and a series of events whose alteration or disruption can result in the breakdown of that process. Standard reproductive toxicity test designs tend to be apical, examining the overall potential for alteration or disruption. They are designed to determine the reproductive competence of the exposed parental generations and the development of their offspring. Generally, both females and males are exposed, and mating is carried out within like exposure groups. The design permits examination of the overall reproductive competence of the parents and maximizes the likelihood of observing a reproductive effect. In cases where a sex-specific effect is expected, a study can be designed to examine this possibility. Standard designs for reproductive toxicity testing do not attempt to identify specific mechanisms of toxicity but rather give general indications of the potential hazard and the types of reproductive parameters that might be sensitive. If appropriate, additional studies can be carried out to explore the specific effects of the agent. It should be recognized that there are a number of effects that usually would not be identified after the standard reproductive study protocols (e.g., transplacental carcinogenesis, reduction of reproductive life spans, and some heritable effects).

The participants discussed a number of end points that are or can be examined under current designs for reproductive effects testing. There was a consensus on the general categories of end points that could be considered as minimum for adequate evaluation of reproductive toxicity, but it was considered inappropriate to detail specific approaches, since (1) different laboratories often use different specific test designs that are nevertheless adequate, and (2) specific regulatory concerns may suggest or require different test designs. The general end point categories that were considered a minimum are listed below and, in some cases, may be obtained from sources other than a reproductive effects study.

Mating—including number of male and female animals mated successfully and time to mating

Conception

Delivery—information on the ability of the dam to deliver, as well as the timing and the product of the delivery

Offspring survival and growth—at least to the time of weaning

Toxicity—general and organ specific

Histopathology—reproductive organs

Details on these end points may be obtained from the Agency's proposed risk assessment guidelines on male and female reproductive toxicity.[2,3] Two additional end points that were identified as possibly making current reproductive toxicity testing more useful are sperm analysis and assessment of estrous cyclicity. The participants questioned requiring such measurements but thought that such data could enhance the interpretation of potential toxicity with little additional resource requirements. For example, spermatid counts could be carried out at the time of male necropsy. Estrous cycling could identify environmental interaction in addition to apparent

nonreceptivity and could be performed when checking for copulatory evidence (e.g., plug or sperm in vaginal smear).

There was general agreement that the data collected should be presented as absolute values. Data evaluation procedures that summarize or condense the database or present values as percentages or indices can cloud or bias the indications provided by absolute values. Absolute values can be converted readily to indices, but it is not always possible to reverse the process, particularly since the procedures involved in calculating indices often are variously defined or ill-defined by individual investigators. This is not intended to discourage summary presentations, since these can be quite useful and can provide insight into the investigator's view of the overall results. Rather, it is intended to stress that such presentations should be additional to, and not a replacement for, presentation of absolute data. If procedures that summarize or condense the data are used, the investigator must be careful that the approach is well-defined and the outcome clearly presented.

The separation of reproductive effects into primary (direct) and secondary (indirect) effects was raised at various points in the discussion of the purpose of a reproductive toxicity study. There was a general lack of support for the separation of effects into primary and secondary. There is not a clear definition of these terms. Moreover, none of the participants could demonstrate a particular role for this separation in the evaluation of reproductive effects testing, since the risk assessment would probably not be affected by whether an effect was primary or secondary.

ISSUE: ADDITIONAL INFORMATION OBTAINED FROM A TWO-GENERATION STUDY

The second issue addressed what additional information is obtained from a two-generation study compared with that from a one-generation study. The workshop used the definitions of the Organization for Economic and Cooperative Development (OECD) for one- and two-generation studies. The number of the generation (e.g., F_1) refers to the filial (F) generation derived from the parental line that was initially exposed (P_0). Exposure of the parental line is initiated in the adult before mating to produce the first generation (F_1) and is continuous to the end of the study (weaning of first- or second-generation offspring) (Fig. 1A).

All of the participants recognized that a one-generation study does not permit evaluation of reproduction in animals exposed prenatally as well as postnatally. It was further pointed out that if meiotically dividing cells are to be considered as targets, evaluation of reproductive performance in females exposed prenatally would be required.

The following summarizes those points that were discussed supporting the use of a two-generation approach. There may be some degree of overlap among these points, and thus, they should not be regarded as being mutually exclusive. Specifically, a two-generation study appears to provide for:

> Evaluation of effects on reproduction where animals have been exposed prenatally as well as postnatally, including the prepubertal period. In other words, a two-generation study allows for the measurement of the maturation of the offspring and the evaluation of organs and systems that are not fully developed at weaning (the time at which a one-generation study is usually terminated).
>
> Evaluation of agents where additional exposure may contribute to the toxicity. The F_1 generation has been potentially exposed indirectly while in utero and via lactation and directly from the time of weaning. Therefore, the F_1 generation has been exposed longer and starting from a younger age than the P_0 generation. In addition, when the inhalation route is used or exposure is by the dietary or drinking water routes, the mg/kg/day dosage is generally increased for the F_1 generation, as compared with the P_0 generation. This additional exposure and increased dosage may be expected to result in increased toxicity.

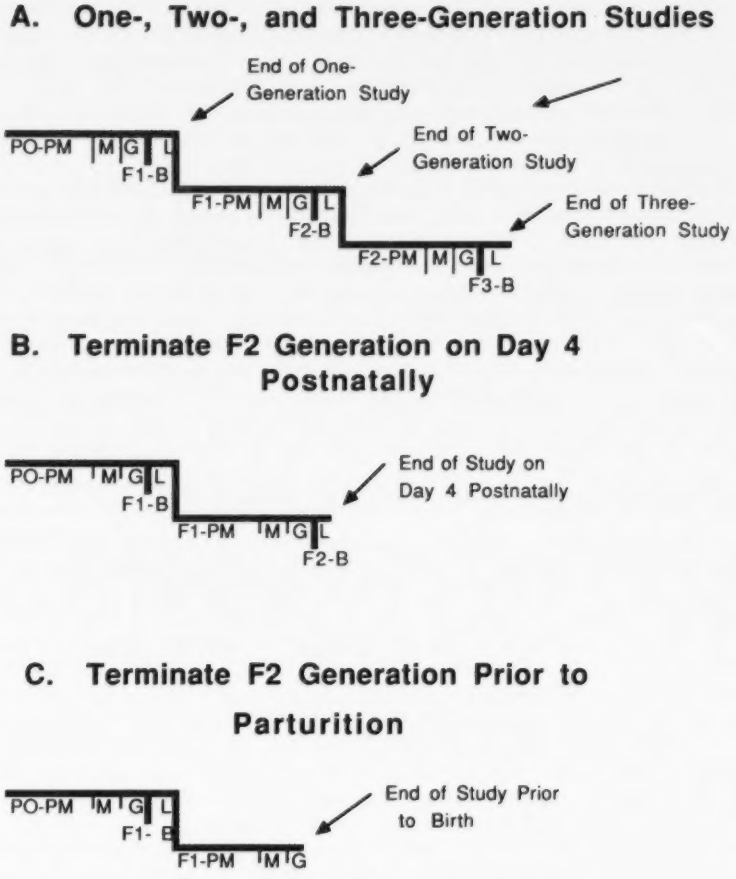

FIG. 1. One-, two-, and three-generation studies and variations. Solid line, exposure period; PO, parental generation; F1, first generation; F2, second generation; F3, third generation; PM, premating period; M, mating period; G, gestation period; L, lactation period; B, birth.

Evaluation of some transmissible germ cell effects. Although a two-generation study should not be regarded as a test to routinely assess all transmissible germ cell effects, it does permit the expression and the detection of some effects that may have a heritable basis.

Evaluation of delayed or latent manifestations of some toxicities. By allowing the F_1 generation to undergo maturation, it may be possible to detect some manifestations that may not be detectable by the time of weaning and the termination of a one-generation study (e.g., gross behavioral alterations, as with exposure to organic mercury and lead). Furthermore, the maturing F_1 generation may be subjected to more specialized testing to detect delayed or latent manifestations of some forms of toxicity (e.g., special measures to assess kidney function or evaluation for the induction of transplacental carcinogenesis, as with diethylstilbestrol).

Detection of effects absent in the first generation, clarification and enhancement of effects that were marginal or not obvious in the first generation, and establishment of effects at lower exposure levels than those that caused toxicity in the first generation.

It should be noted that in saying a two-generation study "provides for" these five evaluations, it does not mean that each of the evaluations is specifically designed into the study. As noted previously, the current two-generation study is an apical test and as such detects, in one or a limited number of end points, the summation of a variety of more specific effects. A detailed analysis of a wide variety of end points generally is not conducted. More often than not, more specific study designs would be required to detail the factor or factors that led to the reproductive toxicity.

Given the limited nature of reports in the literature, it is impossible to determine how often a two-generation reproductive effects study design will provide a more accurate assessment of the potential hazard associated with exposure to a particular agent than a one-generation design. Surveys by Palmer[4] and Clegg[5] that were conducted on unpublished reproductive effects studies performed for regulatory agencies support the idea that a two-generation study is advisable. It was suggested that a review of these more complete databases, in comparison with information found in the open literature, should provide clarification of many of the issues discussed in this workshop (see Considerations for the Future). Currently, the participants thought that two-generation studies were needed to adequately assess an agent's potential as a reproductive toxicant.

ISSUE: ISSUES RELATED TO BIOACCUMULATION

The proposal as presented to the Risk Assessment Forum was phrased: Is a single-generation reproductive effects study sufficient to assess the reproductive toxicity potential of chemicals that do not bioaccumulate? The Panel was asked to address a number of issues related to bioaccumulation. It should be noted that the Panel members were selected based on their expertise in reproductive toxicology and not necessarily in the field of bioaccumulation.

In general, the Panel was reluctant to place much importance on the issue that only agents that bioaccumulate need to be tested in a two-generation study. Reasons for not placing much weight on the issue of bioaccumulation include (1) difficulty in defining "bioaccumulation," (2) difficulty in classifying agents as ones that may or may not bioaccumulate, (3) difficulty in gathering and applying information on an agent's potential to bioaccumulate, (4) limits on the database that is currently available, and (5) most important, that bioaccumulation is not the only critical factor that may account for effects observed in the second generation of animals and not necessarily in the first. The results of the discussions on these issues are expanded below.

The Panel agreed that what constitutes bioaccumulation, in general, is not clearly defined. Consequently, designing studies based on bioaccumulation becomes an impossible task. Often, "bioaccumulation" cannot be distinguished from "accumulation of effects." Both may lead to a progressive decline in function, and both may play a role in reproductive toxicity. In a multigeneration reproductive effects study in which one saw an increasing incidence or magnitude of effects with subsequent generations, it was noted that the following factors could be determined: (1) if the agent itself accumulates in the body, (2) if the effects reflect a cumulative response, or (3) if there truly is a new target or new response of a target (i.e., diethylstilbestrol affected a biological target differently in the F_1 generation than it did in the P_0).[6]

Even if one defines bioaccumulation to mean accumulation within the body or increased body burden with increased duration of exposure, it was considered quite difficult to identify particular chemical classes as prime candidates for consideration as potential bioaccumulating agents. Several Panel members supported the idea that lipid-soluble agents and heavy metals are likely candidates for bioaccumulation, as was proposed by Christian.[1] However, examples were provided to show that not all lipid soluble agents bioaccumulate (e.g., 2,4,5-trichlorophenoxyacetic acid [2,4,5-T] and thiopental), and similarly, that not all metals should be expected to bioaccumulate (e.g., methyl mercury accumulates but ethyl mercury does not). Therefore, much caution was expressed about the dismissal of classes other than those of lipids and metals as not being bioaccumulating agents.

The Panel thought that information on the bioaccumulation of a test agent would be quite valuable before initiation of a reproductive effects study. With such information, it would be possible to combine components of various reproductive toxicity study designs and use scientific knowledge to improve the design on a case-by-case basis. For example, knowing the half-life of an agent may enable one to modify the testing protocol to ensure that animals have reached a steady state by the time mating is to take place. In some cases, such as studies on lead where a steady state cannot be reached until exposure has occurred over quite a long period of time, modification may not be possible. However, the information may help in the interpretation of the results. The fact is that information on bioaccumulation rarely is available. Another example that was suggested was to incorporate a continuous breeding component in those reproductive toxicity studies being conducted on agents known to bioaccumulate. It was emphasized, though, that the designing of a study should not be carried out independently, but should involve discussions with the appropriate regulatory agency to ensure that the study design being developed is not only the most appropriate one scientifically but one that meets the needs of the regulatory agency as well.

Most Panel members believed that it may be more efficient, less costly, and less complex to conduct the second generation in a reproductive effects study than to conduct studies needed to determine the agent's bioaccumulating properties. If studies to assess an agent's ability to bioaccumulate are conducted, they should be conducted in pregnant as well as nonpregnant animals if the information is to be used in conjunction with a reproductive effects study. This is because the metabolism and pharmacokinetics in pregnant and nonpregnant animals may be markedly different.

The Panel agreed with the statement in the Christian review of the literature[1] that the data that were included were extremely limited in nature because of the relatively few published studies (20) identified by the author as adequate for inclusion. The general consensus of the Panel was that these data should be supplemented by review of additional studies. Review of additional data would probably provide better direction as to how to use information on an agent's bioaccumulating properties. It was strongly recommended that data within EPA and FDA files on multigeneration studies and data from the National Toxicology Program on the results of continuous breeding studies be reviewed, since they would provide valuable supplemental information on this and other issues (see Considerations for the Future).

The Panel agreed that the strongest argument against using an agent's bioaccumulating properties as a benchmark to determine whether or not a second generation should be carried forth is that bioaccumulation is not the only critical factor that may account for effects observed in the second generation but not in the first. A number of other critical factors were identified, any of which may require that the F_1 animals be mated and the F_2 offspring be evaluated in order to observe an effect.

First, a more pronounced response in the filial generations is most likely associated with intake of a higher dosage (mg/kg) at an earlier stage of development, particularly when the agent is administered in the diet, drinking water, or by inhalation. The F_1 animals are potentially exposed indirectly in utero and during lactation and directly during the last third of lactation through weaning. In rats, for example, F_1 animals would receive indirect exposure in utero and through lactation and direct exposure from as early as 2 weeks of age; in comparison, exposure of the P_0 generation does not start until the animals are 6–8 weeks of age.

Another critical factor is the difference between developing and adult animals in basal metabolic factors and in their specific ability to detoxify or activate an agent. It is well established that the ability to metabolize compounds is not fully attained in immature animals; for example, in rodents, metabolic capability is not mature until 30–60 days of age.[7] Therefore, animals exposed before this time may be exposed to a different ratio of the parent compound to various metabolites, a different spectrum of metabolites, and over different periods of time. Similarly, detoxification may be retarded in the neonate by slower elimination of potentially toxic agents, since renal maturation is not completed for some period of time postnatally.[8]

Finally, mutational events may occur to the parental germ cell, producing heritable alterations in the offspring. Certain of these alterations may be expressed only when the F_1 generation is mated. At this time, the F_1 male may display semi- or complete sterility. Recent investigations of acrylamide provide a good example of such a transmissible effect.[9]

ISSUE: ALTERNATIVES TO A ONE-GENERATION STUDY

A one-generation reproductive effects study alone was determined to be insufficient to identify all potential reproductive toxicants. The Panel discussed alternatives or optional test designs that would cover the range of effects that they considered important in a reproductive toxicity test design. Most of the suggested alternatives have not been studied thoroughly and merit further attention. They are not viewed currently as replacements for the standard two-generation reproductive effects study. They are offered here as possible alternatives that may be selected on a case-by-case basis depending on what is known about the test agent in question. They are offered to both investigators and regulators as options that may be discussed in negotiations over testing strategies and offer some flexibility in study design. In essence, the use of scientific expertise to arrive at an improved design on a case-by-case basis was encouraged. However, when major diversions from the standard general purpose designs are considered, the testing organization and the appropriate regulatory agency should enter into discussions early to ensure mutual agreement that the strategy is scientifically appropriate.

The alternative methods that were discussed included: (1) methods that incorporate as many evaluations as possible short of being a full-scale two-generation study, (2) alternative methods to a standard one- or two-generation reproductive effects study, most of which include a continuous breeding component, and (3) the use of a three-generation reproductive effects study.

Methods short of being a standard two-generation study

One suggestion was to conduct a modified two-generation study in which the study would be terminated when the F_2 offspring are 4 days old instead of 21 days old (Fig. 1B). This alternative saves 17 days, which may or may not have an economic impact depending on where the study is conducted. The rationale behind this suggestion was that there is little reason to expect any increase in toxicity in the F_2 litters as compared with the F_1 litters during the period of day 4 to day 21. In general, most Panel members thought that saving 17 days was not worth the sacrifice of perhaps missing any effects that may not manifest themselves until after day 4 in the second generation.

Another suggested approach was to start dosing of the parental animals (P_0) at weaning, allow the F_1 animals to mature and mate, terminate the study before parturition, and examine the uterine contents (Fig. 1C). The rationale behind this suggestion was that by starting dosing earlier, earlier periods of the maturation of the reproductive process are covered, and by including a mating component, data are available on the mating and some reproductive outcomes of the F_1 generation. This design would fail to identify any adverse reproductive outcomes occurring after the early killing (i.e., effects on parturition, number of live births, postnatal survival, and growth to weaning of the F_2 generation).

For known or suspected reproductive toxicants, studies of a segmental design, such as those used by regulatory agencies in a number of countries, may be conducted in order to identify which particular aspects of reproduction are affected. The Segment I study is similar in design to a one-generation study, the Segment II study is essentially a standard developmental toxicity or embryo–fetal toxicity/teratogenicity study, and the Segment III study is a peri- and postnatal study. None of these designs, as originally developed[10] (Fig. 2), includes the development of a second generation. Thus, although addressing specific aspects of reproduction more precisely, these segmented studies are more limited in provision of the overall scope of the information they provide than the alternatives previously discussed.

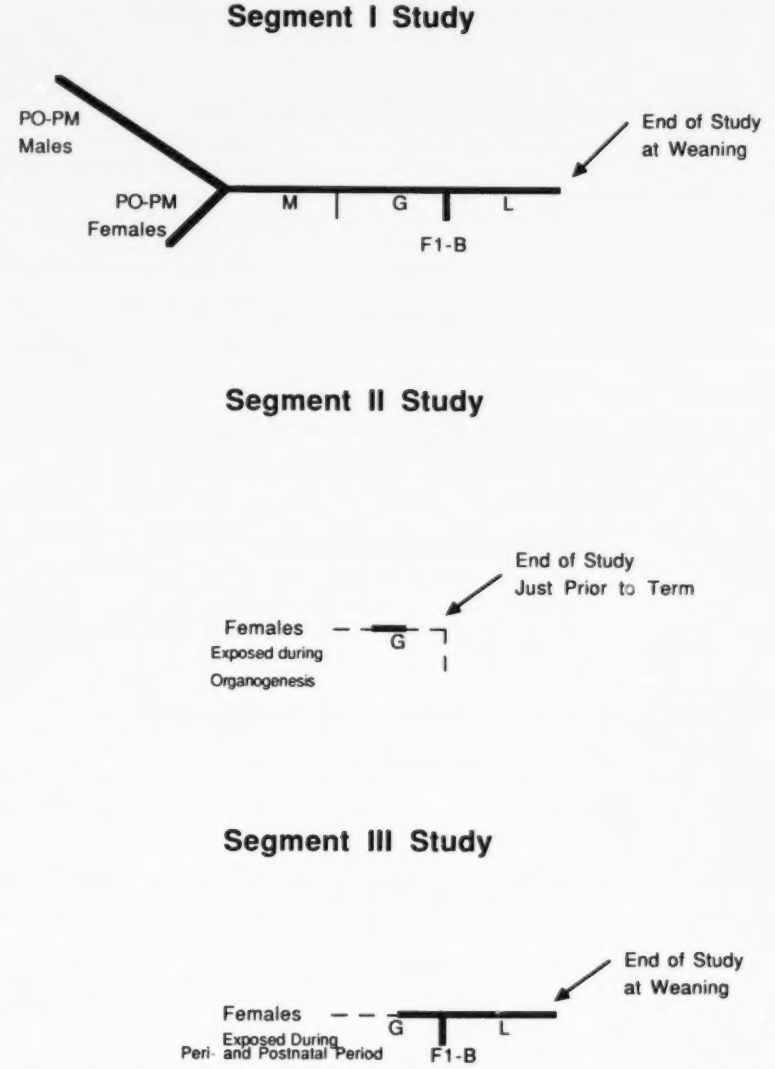

FIG. 2. Segment designs. Solid line, exposure period; broken line, no exposure; PO, parental generation; F1, first generation; PM, premating period; G, gestation period; L, lactation period; B, birth.

Alternative methods to standard one- or two-generation reproductive effects studies

Continuous breeding component. There was considerable discussion about, and considerable support for, the inclusion of a continuous breeding component into a single-generation or modified two-generation study. This was supported by the discussion on bioaccumulation, since

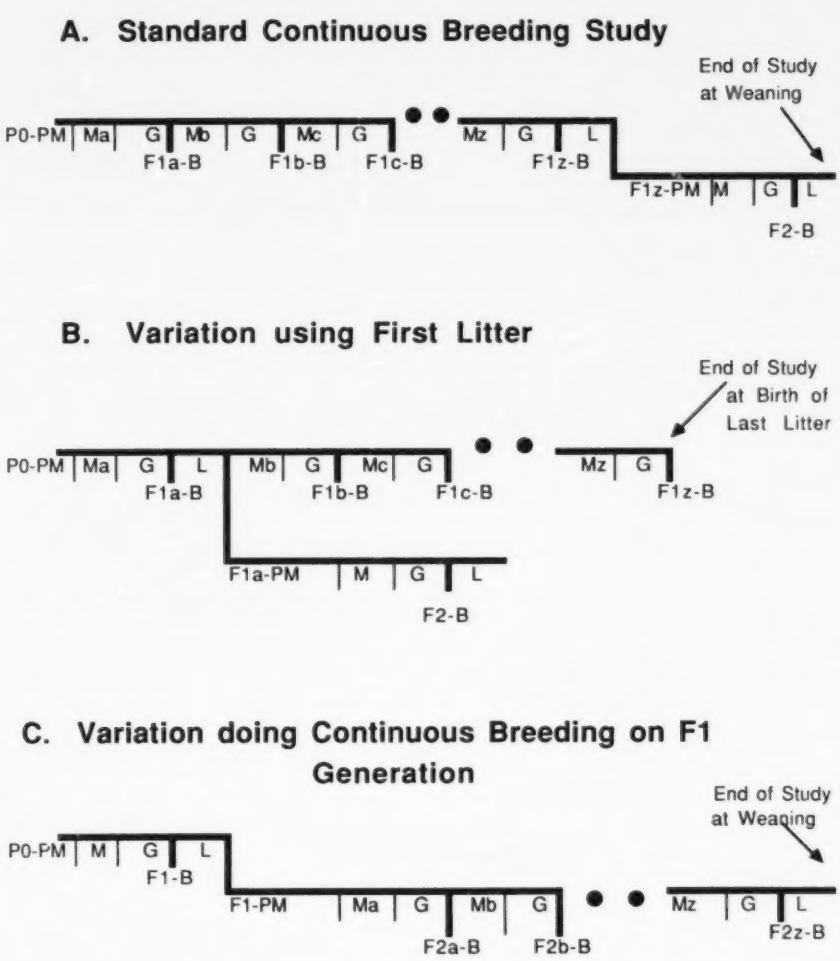

FIG. 3. Continuous breeding studies. Solid line, exposure period; broken line, no exposure; (● ●), mating is repeated until last mating; PO, parental generation; F1, first generation; F2 second generation; PM, premating period; M, mating period; G, gestation period; L, lactation period; B, birth; a, following filial generation, indicates first litter in that generation; z, following filial generation, indicates last litter in that generation.

continuous breeding seems to be one of the most effective ways of investigating bioaccumulation or cumulative effects.

As the continuous breeding study is conducted now, male and female animals are exposed to the test agent for 1 week and are then cohabitated for 98 days while exposure continues, to produce successive litters. As each F_1 litter is delivered, it is examined and killed, and the female is recohabitated with the male to produce another F_1 litter. The last F_1 litter produced is allowed to survive, is dosed from the time of weaning, and is mated to produce a single F_2 litter (Fig. 3A). Various permutations of this design were suggested and discussed.

Rather than use the last F_1 litter to produce the F_2 litter, the first F_2 litter may be used (Fig. 3B). Apart from this variation, the continuous breeding component continues as noted previously. The rationale behind this approach is (1) if there are problems with the first litter, one could use the second, third, and so on litters to produce the next generation; if one waits until the last litter to do so, it is not possible to have a back-up should a problem arise, and (2) the first litter is allowed to mature while the continuous breeding component with the parental animals continues concurrently; this decreases the time it takes to conduct the entire study. For agents that are known or suspected to bioaccumulate, though, it may be more appropriate to use the last litter, since the additional exposure would provide a greater body burden than would occur from using the first litter.

Another variation is to restrict breeding of the P_0 animals to a single litter, confining the role of the P_0 to detecting more immediate effects (e.g., on libido, precoital time, duration of gestation, and neonatal survival) and to providing the F_1 generation (Fig. 3C). Continuous breeding of the F_1 is then conducted either as in the original version or in the first modification. The rationale behind this variation is that since F_1 animals are exposed to the test agent during an earlier period of development, for a longer period, and to higher dosages on a relative body weight basis, potential effects of an agent may be maximized. A caveat is that this may increase the duration of the study and may incorporate more confounding variables.

It was noted that the continuous breeding design results in more complex studies requiring the sacrifice of more animals and may result in difficulty in interpreting toxicity in the parental generation because of interaction of age and body weight on reproductive performance.

Modified two-generation study. Since a shortened period of exposure (1 week) during the premating period was discussed for use in the continuous breeding variations identified previously, the Panel discussed reduction of the period of premating exposure in a standard two-generation study. Currently, the premating period of exposure is 10 weeks. Instead of limiting exposure to 1 week, the animals may be exposed for 4 weeks before mating, as recommended by Palmer[4] in his modified multigeneration approach. This 4-week procedure provides for a longer exposure period that allows additional stages of spermatogenesis to be potentially affected, as well as improving the possibility of identifying an adverse effect. This approach is similar to that depicted in Figure 3B, except that it does not include a continuous breeding component. Although some Panel members were reluctant to accept this alternative without including a continuous breeding component, this method was considered as a potential option that would merit further discussion and evaluation. The rationale behind this approach is to rely more heavily on the results of F_1 maturation and generation of an F_2 litter. The F_1 would serve as the most highly exposed population. Exposure for 1–4 weeks was thought to be sufficient to identify any adverse effects on the libido of the parental generation and to ensure that the animals had some exposure before mating. A perceived limitation of this method and of the continuous breeding component is that the short premating exposure period of the P_0 would allow detection only of effects on late stages of spermatogenesis. Effects on early cell types, such as spermatogonia, would not be detected in view of the long time before their products became mature spermatozoa. As such, some stage-specific male reproductive toxicants may not be identified in the parental generation. However, it is expected that these agents would be identified in the F_1 generation, since these animals (male and female) would be potentially exposed indirectly in utero and via lactation and directly from the time of weaning to the time of mating, so all stages of spermatogenesis and oogenesis would be exposed by the time of the mating of the F_1 generation.

Three-generation reproductive effects study

In addition to considering the mentioned alternatives, the Panel also considered the necessity of a three-generation reproductive effects study (Fig. 1A). Protocols (both in the United States and in Japan) allow for the production of a third generation of animals in certain cases, and the Panel agreed that this approach should not be ruled out entirely. The Panel recognized that the data may be particularly useful in claryifying findings of the first and second generations.

However, several of the participants pointed out that three or more generations often confuse interpretation of the results, especially if adverse effects are observed in some generations and not in others (e.g., in only the second generation). Thus, the participants concluded that, in general, carrying a study beyond the second generation is unsupported, may be confusing, and should be used only on a case-by-case basis.

Additional Comments

Regardless of the methodology used to assess reproductive toxicity, the Panel thought that it is important to supplement the results of any such study by reviewing and considering the results of all toxicity studies, such as the dominant lethal, developmental toxicity, Chernoff–Kavlock, and subchronic and chronic studies (specifically the histopathology of the reproductive organs). This relates to the Panel's philosophy regarding the purpose of a reproductive toxicity study, which was presented earlier and which referred to assessment of a reproductive hazard in context with other forms of toxicity. However, there should still be a concern for any adverse reproductive effects, regardless of whether or not there are other forms of toxicity present.[2,3]

The Panel believes that the time has come to allow the use of scientific judgment in test design. It is important to realize that there are many different tools available to reproductive toxicologists. The indicated alternative methodologies are offered by the Panel as recommendations that may be selected for use in testing, depending on a particular need or question raised. However, any protocol employed should still provide the end points described previously as minimum end points necessary to assess reproductive toxicity. It was recommended that the EPA consider the development of additional testing protocols in the future to add to the limited list of protocols that it currently uses. This would afford greater flexibility and variety in choosing a protocol on a case-by-case basis.

CONSIDERATIONS FOR THE FUTURE

After discussing the issues noted, the workshop participants considered several areas where additional work may provide information to fill gaps that currently exist.

Review of databases of regulatory agencies and programs

An issue addressed in the Christian review[1] and also recognized in the preparation phase for and during the workshop was the limited number of well-conducted or well-reported one- or two-generation reproductive effects studies in the open literature. Not only are well-conducted and well-reported studies difficult to identify, but also reproductive toxicity studies that demonstrate no reproductive or other forms of toxicity are often not published, thus biasing the review to those studies that have some positive findings. Thus, an attempt to address data gaps using the open literature can be expected to be unproductive or, at best, of limited use.

A suggestion was made to access the databases that could be made available from regulatory agencies (e.g., EPA and FDA), testing programs (National Toxicology Program), or other contributors (e.g., interested industrial groups) and to review these studies relative to the indicated issues as well as to alternatives in study design and data presentation. It appears that these databases could provide as many as 100–150 well-conducted studies and that the agents tested would cover a wide range of physical-chemical characteristics and intended uses and exposures. It is recognized that many of these data are proprietary; therefore, questions would have to be designed and data reviewed in such a manner as to assure the confidentiality of the information. This may preclude publicly reporting specific results on certain issues.

It must be recognized that such a review will require a considerable resource commitment in terms of personnel and funding. Such reviews cannot be carried out quickly and will require detailed review of many end points in order to establish the presence or absence of toxicity. It

will also require reviewers who are knowledgeable in the areas of reproductive and developmental toxicity and experienced in the conduct and evaluation of such studies.

Major symposium–workshop

An approach that the Agency has used in the past for addressing issues of concern to the field of reproductive and developmental toxicity is to hold a large workshop of experts to address specific questions that can be derived from the current state of our knowledge. The general feeling of the Panel was that a larger workshop would be premature and should await a review of larger databases, as suggested previously. In addition, any larger workshop would best be held after the publication and public comment period of the Agency's proposed guidelines on the risk assessment of male and female reproductive toxicity. In this way, the issues and questions would be better defined and would permit a more substantive response from the participants. It is appropriate for the publication of the present workshop to be open for public comment, allowing for input from the broader scientific and regulatory community.

Research needs

The participants agreed that the review of internal databases, as previously described, was a primary research need. In addition, there was a discussion of research needs as they relate to alternative test systems. Some participants expressed the opinion that test designs that were specific alterations of the one- or two-generation design would have to be evaluated for their ability to define the reproductive toxicity of an agent. Other participants thought that such evaluation would be difficult and questioned whether enough data were available on a large enough group of test agents to permit adequate review. All participants agreed that review of the databases from regulatory agencies and programs, as discussed previously, would be the most important approach to addressing a number of issues related to reproductive toxicity testing. As this review is carried out, research needs will become more obvious.

Reporting of reproductive effects studies

The large volume of data generated by a reproductive effects study raises the issue of how best to report it. In the past, there has not been an established format for such reporting. The point was made that such a format may assist in the review process and that a minimum set of end points (e.g., those end points discussed under the first issue) should be established to assure the adequacy of the database for review. As previously discussed, the reporting of individual data points was considered axiomatic. However, a rigid framework for data analysis and reporting can stifle the investigator's ability to accurately describe the findings, should not be imposed without good reason, and may be counterproductive.

A point made throughout the workshop was that reporting of reproductive effects studies in the peer-reviewed literature is often inadequate for regulatory purposes. Because such reports are often integral to the risk assessment process, it is imperative that journal editors and reviewers require that articles include a minimum set of end points and an accurate description of the reproductive findings relative to an agent's potential toxicity. It is recognized that space restrictions will limit data presentation to summary values, but either the journal or the author should have more complete information on procedures and data that can be made available to those who request it.

DISCUSSION

The workshop was essentially organized to address the central question: Is a single-generation reproductive effects study sufficient to assess the reproductive toxicity potential of chemicals that

do not bioaccumulate? Although this question served as a starting point, the workshop went beyond and considered some broader issues as well. In response to the question, the Panel concluded that, by itself, a one-generation reproductive effects study is insufficient to identify all potential reproductive toxicants, and that, currently, a two-generation study is needed for an adequate assessment. Use of a two-generation approach was supported by a number of points that illustrated what additional information could be obtained from a two-generation study as compared with a one-generation study. Furthermore, the Panel did not support placing much importance on bioaccumulation as related to determining the appropriateness of one- or two-generation reproductive toxicity testing. A number of technical factors, such as difficulty in defining and testing for bioaccumulation, were discussed as confounders. However, the main reason for not placing much weight on bioaccumulation is that it is not the only critical factor that may account for effects in a second generation but not in the first. The Panel strongly believed that information on the bioaccumulation of a test agent would be highly useful before the start of a reproductive effects study, since it would allow for adjustment of the study design on a case-by-case basis.

In looking at several broader issues, the Panel accomplished several goals that hopefully will serve to guide or direct future development in reproductive toxicity testing. First, a clearly defined purpose of a reproductive effects study, in the context of risk assessment, was developed and served to guide subsequent discussions. In addition, a minimum set of end points necessary for adequate evaluation of reproductive toxicity was determined. The Panel made recommendations of a potential list of some alternative reproductive effects test methods that may be considered in the development of testing strategies. The participants encourage the use of scientific judgment in designing test methodologies and, therefore, recommend the development of additional testing guidelines by regulatory agencies to permit greater flexibility in choosing a test protocol on a chemical-by-chemical basis.

Areas in which additional work is needed to permit more complete assessments of reproductive toxicity were identified by the Panel. The project that was considered of greatest priority is a review of databases from regulatory agencies and programs. The information from such a review will be highly useful in designing better test methodologies, identifying research needs, and identifying outstanding issues in reproductive effects testing and assessment of results that may best be addressed at a major symposium or workshop.

This workshop opened up a dialogue among reproductive scientists within regulatory agencies and testing laboratories to study issues of reproductive effects testing in the context of the scientific state-of-the-art and the current regulatory framework. It should be recognized that the workshop focused on specific issues related to the central question and that there are many other issues related to reproductive toxicity testing, in general, that were not discussed. Therefore, it is hoped that this dialogue will continue and will stimulate progress in this field not only in terms of laboratory investigations but also in developing improved study designs, increasing the quality of data reporting, and conducting database reviews.

ACKNOWLEDGMENT

The organizers are indebted to the participants for their careful preparation and insightful comments during the workshop and the editing of the publication. We would also like to thank EPA's Risk Assessment Forum for their support and Eastern Research Group for their assistance in organizing the workshop and the final report.

REFERENCES

1. CHRISTIAN, M.S. (1986). A review of multigeneration studies. J. Am. Coll. Toxicol., **5**, 161–180.

2. U.S. ENVIRONMENTAL PROTECTION AGENCY. (1988a). Proposed guidelines for assessing male reproductive risk. Fed. Register **53,** 24850–24869.
3. U.S. ENVIRONMENTAL PROTECTION AGENCY. (1988b). Proposed guidelines for assessing female reproductive risk. Fed. Register **53,** 24834–24847.
4. PALMER, A.K. (1986). A simpler multigeneration study. Poster presented to 6th IUPAC Congress of Pesticide Chemistry, Ottawa, Canada.
5. CLEGG, D.J. (1979). Animal reproduction and carcinogenicity studies in relation to human safety evaluation. In: *Toxicology and Occupational Medicine.* Diechman, W.B. (ed.). Amsterdam: Elsevier, pp. 44–59.
6. LAMB, J.C., JAMESON, C.W., CHOUDRY, H., and GULATI, D.K. (1985). Fertility assessment by continuous breeding: Evaluation of diethylstilbestrol and a comparison of results from two laboratories. J. Am. Coll. Toxicol. **4,** 173–184.
7. PARKE, D.V. (1984). Development of detoxification mechanisms in the neonate. In: *Toxicology and the Newborn.* Kacew, S., and Reasor, M.J. (ed.). Amsterdam: Elsevier, pp. 1–31.
8. KAVLOCK, R.J., and GRAY, J.A. (1982). Evaluation of renal function in neonatal rats. Biol. Neonate **41,** 279–288.
9. SHELBY, M.D., CAIN, K.T., CORNETT, C.V., and GENEROSO, W.M. (1987). Acrylamide: Induction of heritable translocation in male mice. Environ. Mutagen. **9,** 363–368.
10. FOOD AND DRUG ADMINISTRATION. (1966). Guidelines for reproduction and teratology testing of drugs. Bureau of Foods.

Address reprint requests to:
Elaine Z. Francis
U.S. Environmental Protection Agency
(RD-689)
401 M Street, S.W.
Washington, DC 20460